

Impact of generative artificial intelligence-based virtual tutors on self-directed mathematics learning: a systematic review

Carlos G. Portilla Peralta 

Universidad Nacional de San Agustín, Arequipa, Perú, cportillap@unsa.edu.pe

ABSTRACT - *The emergence of Generative Artificial Intelligence (GenAI) and Large Language Models (LLMs) reconfigures mathematics education, shifting the boundaries of traditional instruction toward dynamic conversational tutoring models. Objective: This study analyzes the impact of GenAI-based virtual tutors on self-directed learning in mathematics, identifying pedagogical trends, benefits for autonomy, and implementation challenges. Methodology: The authors conducted a systematic literature review of studies published between 2022 and 2026, following PRISMA guidelines. They restricted the search to Spanish-language publications to specifically analyze the technological impact within the Ibero-American and Spanish-speaking educational context. Furthermore, they evaluated the methodological rigor of the selected studies using standard tools (CASP and statistical criteria). Results: The highest-quality methodological studies indicate that GenAI integration significantly improves academic performance and the understanding of abstract concepts (such as functions and calculus), while reducing math anxiety and cognitive load. These tools act as scaffolding that fosters self-regulation through immediate and personalized feedback. However, researchers identified critical risks associated with data accuracy ("hallucinations") and a potential technological dependence that limits critical thinking. Conclusions: Virtual tutors are effective catalysts for the personalization of mathematical learning, provided their implementation is mediated by critical digital literacy strategies and instructional design that prioritizes human validation. Finally, due to the language restriction, these findings reflect the specific realities of the Spanish-speaking educational culture and are not directly generalizable worldwide.*

Keywords: *Generative Artificial Intelligence, virtual tutors, self-directed learning, mathematics education, systematic review.*

I. INTRODUCCIÓN

La integración de la Inteligencia Artificial (IA) en el ámbito educativo ha experimentado una transformación radical en los últimos años, desdibujando las fronteras tradicionales entre la instrucción humana y el soporte tecnológico [1]. En particular, la irrupción de la Inteligencia Artificial Generativa (IAG) y los Modelos de Lenguaje Grande (LLM) ha generado un nuevo paradigma en la enseñanza superior, donde herramientas como

ChatGPT no solo funcionan como repositorios de información, sino como "socios activos" capaces de revolucionar los métodos de evaluación y el apoyo al estudiante en entornos complejos [2]. Este avance tecnológico plantea interrogantes críticos sobre cómo estas herramientas pueden moldear las competencias de los ciudadanos del mañana, especialmente en disciplinas fundamentales caracterizadas por altos niveles de abstracción, como las matemáticas.

En el contexto de la educación matemática, los Tutores Inteligentes (ITS) y los modelos adaptativos se han posicionado como una solución prometedora para personalizar el aprendizaje. A diferencia de los métodos tradicionales, los sistemas actuales impulsados por IA permiten adaptar los contenidos al razonamiento algebraico y a las necesidades específicas de cada alumno, facilitando una comprensión más profunda de temas complejos como las funciones en el bachillerato y el cálculo multivariado a nivel universitario [3], [4]. La literatura reciente sugiere que la integración de estas herramientas tiene el potencial de mejorar la retención del conocimiento y reducir las tasas de deserción mediante intervenciones personalizadas que analizan grandes volúmenes de datos educativos con alta precisión [1], [5].

Sin embargo, el verdadero valor de estas tecnologías no reside únicamente en la transmisión de contenidos, sino en su capacidad para fomentar el aprendizaje autodirigido y la autorregulación. Estudios recientes destacan la importancia del aprendizaje autorregulado en entornos virtuales universitarios, donde la IA actúa como un andamiaje que permite al estudiante gestionar su propio ritmo, reducir la carga cognitiva extrínseca y recibir retroalimentación inmediata [6], [7]. La eficiencia de los asistentes virtuales y la calidad de su retroalimentación la cual puede estructurarse bajo modelos pedagógicos como la "Escalera de Wilson" han demostrado ser factores determinantes en la satisfacción estudiantil y en el desarrollo de la autonomía académica [8], [9].

A pesar del entusiasmo por la implementación de chatbots y asistentes generativos, persisten debates significativos sobre sus implicaciones éticas y pedagógicas. El uso de estas estrategias conlleva desafíos relacionados con la dependencia tecnológica, la falta de precisión o "alucinaciones" de los modelos, y la necesidad urgente de una alfabetización digital que incluya la ingeniería de *prompts* [10], [4]. Es imperativo analizar no solo la dimensión tecnológica, sino también los factores sociales y educativos que influyen en su implementación efectiva para evitar exacerbar las desigualdades existentes.

En consecuencia, la presente revisión sistemática se plantea para esclarecer el estado actual de esta disrupción

tecnológica en el periodo 2022-2026, estableciendo los siguientes objetivos de investigación:

Objetivo General

Analizar el impacto de los tutores virtuales basados en Inteligencia Artificial Generativa en el aprendizaje autodirigido de las matemáticas, mediante una revisión sistemática de la literatura científica publicada entre 2022 y 2026, para identificar las principales tendencias pedagógicas, beneficios en la autonomía del estudiante y desafíos en su implementación.

Objetivos Específicos:

- Identificar las estrategias didácticas y tendencias pedagógicas emergentes que integran la Inteligencia Artificial Generativa en la educación matemática, tales como la tutoría automatizada, el diseño de *prompts* y la retroalimentación formativa personalizada.
- Evaluar el impacto de los tutores virtuales en el desarrollo de la autonomía, la autorregulación y el rendimiento académico, analizando la evidencia sobre la mejora en la resolución de problemas, la gestión del ritmo de aprendizaje y la reducción de la ansiedad matemática.
- Analizar los desafíos éticos, técnicos y de implementación asociados al uso de la Inteligencia Artificial en el contexto educativo, examinando barreras como las "alucinaciones" algorítmicas, el riesgo de dependencia tecnológica y la necesidad de validación crítica por parte del alumnado.

II. METODOLOGÍA

El presente estudio se desarrolló bajo un diseño de revisión sistemática de la literatura, siguiendo las directrices de la declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). El objetivo fue analizar la producción científica sobre el impacto de la Inteligencia Artificial Generativa y los tutores virtuales en el aprendizaje autodirigido de las matemáticas.

A. Estrategia de Búsqueda

La búsqueda exhaustiva se realizó en las bases de datos Scopus y Google Scholar, cubriendo el periodo de publicación comprendido entre 2022 y 2026. Se diseñó y ejecutó una cadena de búsqueda booleana estructurada para recuperar literatura que interceptara tres ejes temáticos: tecnologías de IA generativa, educación matemática y aprendizaje autónomo.

Fórmula de búsqueda utilizada: ("Generative AI" OR "GenAI" OR "Large Language Models" OR "LLM" OR "Virtual Tutor*" OR "Intelligent Tutoring System*" OR "Chatbot*") AND ("Math*" OR "Mathematics education") AND ("Self-directed learning" OR "Self-regulated learning" OR "student autonomy")

B. Criterios de Selección

TABLA I
CRITERIOS DE INCLUSIÓN/EXCLUSIÓN

Criterios de inclusión	Criterios de exclusión
Población: Estudiantes de educación secundaria y universitario (especialmente en el área de matemáticas).	Población: Estudiantes de niveles iniciales, universitarios o entornos profesionales.
Intervención: Uso de tutores virtuales impulsados por IA Generativa o modelos de lenguaje.	Intervención: Herramientas de software educativo estático o sin interacción inteligente.
Temporalidad: Artículos publicados entre 2022 y la actualidad (2026).	Temporalidad: Publicaciones anteriores al año 2022.

Idioma: Investigaciones publicadas originalmente en español.	Idioma: Artículos en inglés u otros idiomas.
Tipo de estudio: Investigaciones empíricas, artículos de revistas indexadas y reportes técnicos de implementación pedagógica.	Tipo de estudio: Literatura gris, presentaciones de diapositivas, blogs o material no arbitrado.

C. Proceso de Cribado

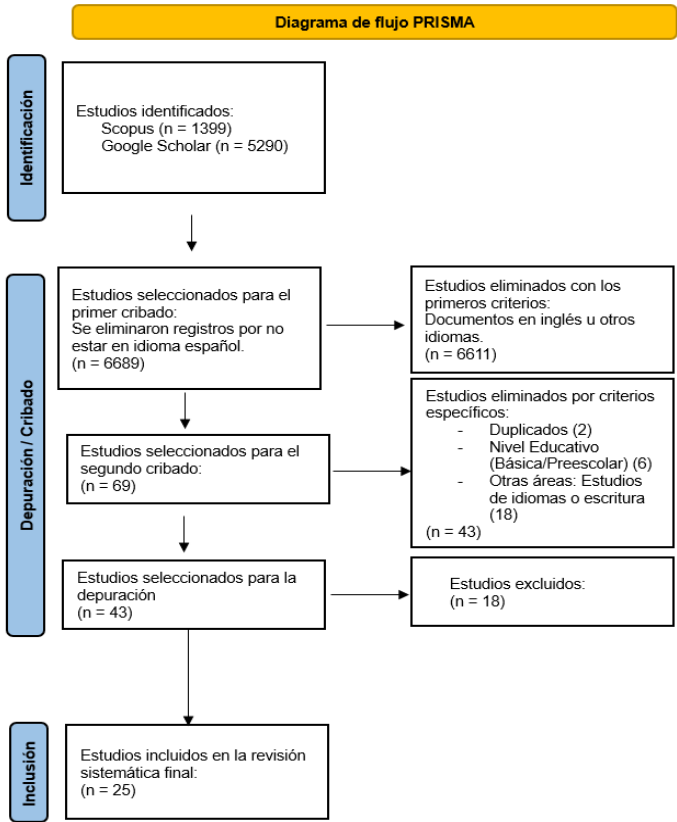


Fig. 1 Diagrama de flujo PRISMA

Los investigadores consultaron las bases de datos Scopus y Google Scholar para el período 2022-2026 en la fase de identificación inicial encontraron un total de 6 689 registros provenientes de Scopus (1 399) y Google Scholar (5 290). Tras aplicar un primer cribado basado en el idioma, seleccionando exclusivamente estudios publicados en español. Esta decisión metodológica busca analizar específicamente el impacto, la adopción y las barreras de la IAG en el contexto educativo hispanohablante e iberoamericano, evaluando cómo estas tecnologías se integran en una cultura educativa con características y brechas de infraestructura particulares, de los cuales eliminaron 6 611 documentos por no estar publicados en español, se seleccionaron 69 estudios potenciales para una evaluación detallada.

Posteriormente, se realizó un segundo cribado empleando criterios específicos de elegibilidad, lo que permitió excluir registros duplicados (2), estudios centrados en niveles educativos básicos o preescolares (6) e investigaciones de áreas impertinentes como idiomas o escritura (18), resultando en 43 artículos seleccionados para la etapa de depuración.

En la fase final de depuración, se excluyeron 18 documentos adicionales tras un análisis exhaustivo de su contenido y metodología, lo que derivó en una muestra definitiva de 25 artículos incluidos en la revisión sistemática. Este proceso de selección, representado en el diagrama PRISMA (Fig. 1), garantiza

que la síntesis cualitativa se fundamente en estudios que cumplen estrictamente con los criterios de calidad y relevancia disciplinar definidos para esta investigación.

D. Evaluación formal de la calidad metodológica

Para garantizar la confiabilidad de la evidencia, los investigadores evaluaron el rigor metodológico de los 25 estudios incluidos utilizando la lista de verificación CASP (Critical Appraisal Skills Programme) para investigaciones cualitativas y parámetros de validez estadística para los estudios cuantitativos (tamaño de la muestra, claridad de objetivos y solidez de conclusiones). Los hallazgos se clasificaron en calidad Alta, Moderada o Baja, como resume la Tabla V.

III. RESULTADOS

A. Características metodológicas y calidad de la evidencia

TABLA II
SÍNTESIS METODOLÓGICA, IMPACTO DE LA IA EN LA EDUCACIÓN MATEMÁTICA Y SUPERIOR

Referencia (Autor y Año)	Tipo de Estudio / Diseño	Enfoque Metodológico	Muestra / Corpus	Nivel Educativo	Herramienta Tecnológica
Estudios Empíricos (Experimentales, Cuasi-experimentales, Correlacionales y Casos)					
Peña-Aguilar et al. [21]	Correlacional / Descriptivo	Cuantitativo	n=569 estudiantes	Educación Superior	Plataformas digitales impulsadas por IA
Cabeza-Rodríguez [8]	Estudio de Caso / Descriptivo	Mixto	n=391 estudiantes	Educación Superior (Online)	Asistente virtual basado en ChatGPT 3.5
Delgado Santin et al. [3]	Cuasi-experimental	Mixto	n=240 estudiantes	Secundaria (Bachillerato)	Modelo adaptativo tipo GPT-4
Guisheca Ayala et al. [15]	Experimental (Pretest-Posttest)	Cuantitativo	n=120 estudiantes	Educación Básica	Plataformas de IA personalizadas
Bonilla M. et al. [4]	Investigación Acción Participativa	Mixto	n=66 estudiantes (3 grupos)	Educación Superior (Ingeniería)	ChatGPT, Gemini, Copilot, GeoGebra
Plaza García et al. [5]	Cuasi-experimental	Cuantitativo	n=28 estudiantes	Secundaria (Bachillerato)	ChatGPT
Marcillo et al. [23]	Estudio de Caso	Cuantitativo	n=25 estudiantes	Educación Superior	App Pyto (Programación móvil)
Leiva-Guerrero et al. [9]	Exploratorio / Validación	Cualitativo	n=2 cursos (validación n=8 expertos)	Educación Superior	GPT-4o (Escalera de Wilson)
Puertas y Cano [24]	Experimental (Prompting)	Cuantitativo / Cualitativo	n=1173 fragmentos de texto	Educación Superior	GPT-4 Turbo (Zero-shot / Few-shot)
Calla et al. [25]	Correlacional / Explicativo	Cuantitativo	Estudiantes universitarios*	Educación Superior	Inteligencia Artificial (General)
Revisiones Sistemáticas, Bibliométricas y Teóricas					
Reyes Vallejo et al. [1]	Revisión Sistemática / Bibliométrica	Cuantitativo	n=413 artículos	Educación Superior	Algoritmos matemáticos de IA
Gutiérrez-Castillo et al. [12]	Revisión Sistemática (PRISMA)	Mixto	n=178 artículos	Educación Superior	IA, Machine Learning, Deep Learning

Reyes-Moreno et al. [13]	Análisis Bibliométrico	Mixto	n=92 documentos	General / Superior	ChatGPT, Chatbots
Vera-Fernández y Sánchez [11]	Revisión Sistemática (PRISMA/PI CO)	Cuantitativo	n=63 artículos	Instituciones Educativas	Chatbots Generativos (ChatGPT)
Torres Valladares et al. [6]	Revisión Sistemática (PRISMA-P)	Cualitativo	n=45 artículos	Educación Superior	Entornos Virtuales / E-learning
Fernández Otoya et al. [14]	Revisión de Literatura (Meta-síntesis)	Cualitativo	n=40 documentos	Educación Superior (Prácticum)	IA Generativa
Maguiña Huerta y Padilla [16]	Revisión Sistemática (PRISMA)	Cualitativo	n=31 artículos	Educación Básica	Apps digitales, Gamificación
Reina-Parrado et al. [17]	Revisión Sistemática (PRISMA)	Mixto	n=25 artículos	General / Educación	Machine Learning (ML)
Díaz Basurco [10]	Revisión Crítica / Tesis	Cualitativo	n=24 investigaciones	Educación Superior (Matemática)	IA Generativa
González Murillo et al. [18]	Revisión de Alcance (Scoping Review)	Cualitativo	n=20 estudios	Educación Superior (Matemática)	ChatGPT
Rodríguez Cairo y Ramírez [19]	Revisión Bibliográfica	Cualitativo	n=15 artículos	Educación Superior	Tutores Inteligentes (ITS)
Reyes Contreras et al. [20]	Revisión Sistemática (PRISMA)	Cualitativo	n=14 artículos	Superior / Secundaria	Tutoría entre pares (contexto IA)
Pezzini [26]	Revisión y Análisis	Cualitativo	n=10 Sistemas (STI)	Secundaria	Sistemas Tutores Inteligentes (Aleks, etc.)
Guàrdia et al. [2]	Editorial / Ensayo Teórico	Cualitativo	N/A (Análisis conceptual)	General	IA Generativa, ChatGPT, LLM
Sorbara [22]	Ensayo Teórico	Cualitativo	N/A (Reflexión)	General	Inteligencia Artificial

El análisis de los artículos seleccionados, sistematizados en la Tabla 2, revela una diversidad metodológica que abarca desde diseños experimentales rigurosos hasta revisiones sistemáticas basadas en protocolos estandarizados. Esta heterogeneidad permite abordar el fenómeno de la Inteligencia Artificial (IA) en la educación matemática y superior desde múltiples dimensiones: el rendimiento académico, la percepción estudiantil y la evolución de la producción científica.

Diseño y enfoque de la investigación: En la literatura analizada se identifica una tendencia hacia enfoques cuantitativos y mixtos, orientados a medir el impacto directo de las intervenciones tecnológicas. Destacan los estudios de corte experimental y cuasi-experimental que emplean diseños de pre-test y post-test para evaluar la ganancia en el aprendizaje de competencias matemáticas específicas, como el razonamiento algebraico y el análisis de funciones [3], [5]. Asimismo, son relevantes las investigaciones correlacionales y descriptivas que buscan establecer la relación entre el uso de plataformas inteligentes y variables como la satisfacción o el rendimiento académico, utilizando instrumentos validados estadísticamente como el cuestionario COMUNICA [8], [21].

Por otro lado, las revisiones sistemáticas y bibliométricas constituyen una parte sustancial del corpus, mostrando una fuerte adherencia a la declaración PRISMA para garantizar la transparencia en la selección de la evidencia [12], [17]. Estas revisiones suelen apoyarse en herramientas de software como VOSviewer o Bibliometrix para mapear redes de coautoría y tendencias temáticas, evidenciando un crecimiento exponencial de la producción científica en el periodo pospandemia [1], [13].

Muestra y contexto educativo: El contexto predominante en los estudios es la Educación Superior, reflejando el interés estratégico por integrar la IA en la formación profesional universitaria y la gestión administrativa [14], [11]. Los tamaños de muestra varían significativamente según el diseño metodológico: desde estudios de caso focalizados con grupos pequeños, como los 25 estudiantes de ingeniería que utilizaron aplicaciones móviles para matemáticas discretas [23], hasta estudios cuantitativos robustos con muestras censales superiores a los 500 participantes [21].

No obstante, la educación básica y el bachillerato también están representados, especialmente en investigaciones que prueban plataformas adaptativas y gamificación para reducir brechas en habilidades fundamentales [15], [16]. Este espectro educativo demuestra que la integración de la IA no es exclusiva de un nivel formativo, aunque su implementación y complejidad varían según la madurez cognitiva del estudiante.

En conclusión, la evidencia recabada demuestra una transición en la investigación educativa sobre IA se ha pasado de estudios meramente técnicos o descriptivos a investigaciones más complejas que evalúan la eficacia pedagógica y la experiencia del usuario. Los resultados sugieren que, si bien la infraestructura tecnológica y el acceso a herramientas son altos, estos no garantizan por sí solos una mejora en el rendimiento académico. El factor determinante parece residir en la mediación pedagógica, el diseño instruccional y la capacidad de las herramientas para ofrecer una personalización real del aprendizaje, superando el uso instrumental para fomentar competencias de orden superior como el pensamiento crítico y la autorregulación.

B. Análisis de los estudios empíricos: Metodología, rendimiento e impacto en la educación matemática

TABLA III
CARACTERÍSTICAS METODOLÓGICAS DE LOS ESTUDIOS EMPÍRICOS INCLUIDOS

Referencia	Etapas Educativas	Tipo de Diseño	Enfoque Metodológico	Población / Muestra
Ramírez-Medina et al. [21]	Educación Superior (Pregrado)	No experimental / Correlacional	Cuantitativo	$n = 569$ estudiantes de Administración (censo).
Cabeza-Rodríguez [8]	Educación Superior (En línea)	Estudio de Caso / Descriptivo e Inferencial	Mixto (predominio cuantitativo)	$n = 391$ estudiantes de 5 centros académicos.
Delgado Santin et al. [3]	Educación Secundaria	Cuasi-experimental (Pretest-Posttest)	Mixto	$n = 240$ estudiantes de tres colegios públicos.
Leiva-Guerrero et al. [9]	Educación Superior	Exploratorio / Validación (Método Delphi)	Cualitativo	$n = 144$ estudiantes (dos cursos) y 8 expertos.
Plaza García et al. [5]	Educación Secundaria (Bachillerato)	Experimental (Grupo Control y Experimental)	Cuantitativo	$n = 28$ estudiantes (14 control, 14 experimental).
Calla et al. [25]	Educación Superior	No experimental /	Cuantitativo	Estudiantes universitarios

		Correlacional-Explicativo		(Piloto $n = 50$).
Marcillo y Carrasco [23]	Educación Superior (Ingeniería)	Estudio de Caso	Cuantitativo	$n = 25$ estudiantes de Ingeniería en Sistemas.
Puertas Cano [24]	Educación Superior	Experimental (Validación de prompts)	Mixto	Fragmentos de retroalimentación en tareas iterativas.

El análisis de los estudios empíricos seleccionados, sistematizados en la Tabla 3, permite identificar patrones significativos respecto a la distribución por nivel educativo, el rigor metodológico y la efectividad de las intervenciones con Inteligencia Artificial (IA) en el aprendizaje. Se observa una predominancia de investigaciones en el nivel de Educación Superior, aunque emergen hallazgos relevantes en la educación secundaria que sugieren una aplicabilidad transversal de estas tecnologías.

Características metodológicas y tamaño del efecto: La evidencia empírica se sustenta principalmente en diseños cuantitativos y mixtos con muestras variables. En el ámbito universitario, destacan estudios con muestras robustas, como el de Ramírez-Medina et al. [21], quienes analizaron a 569 estudiantes para correlacionar la accesibilidad de la IA con la planificación académica, y el de Calla et al. [25], que examinaron las actitudes de más de 300 estudiantes hacia la IA y su relación con la satisfacción académica ($R^2 = .527$). En contraste, las intervenciones en educación matemática específica, como el uso de la app Pyto para matemáticas discretas [23] o el uso de ChatGPT para funciones [5], tienden a utilizar muestras más reducidas ($n < 30$), lo que permite un control experimental más estricto, pero limita la generalización masiva de los resultados.

Impacto en el rendimiento académico y competencias matemáticas: Los resultados cuantitativos indican consistentemente que la integración estructurada de la IA mejora el rendimiento. Plaza García et al. [5] reportaron diferencias estadísticamente significativas ($p < .001$) entre grupos, donde los estudiantes que utilizaron ChatGPT para aprender funciones, dominio y recorrido obtuvieron una media de 7.75 frente a 5.79 del grupo control. De manera similar, Delgado Santin et al. [3] encontraron que el uso de un modelo adaptativo en secundaria generó una mejora del 33.1% en el razonamiento algebraico, superando el 10.2% del grupo control, con un tamaño del efecto grande ($d = 1.4$). Estos hallazgos sugieren que la IA es particularmente eficaz para personalizar el aprendizaje en áreas de alta abstracción matemática.

Satisfacción, actitudes y retroalimentación: Más allá del rendimiento numérico, los estudios abordan la experiencia del estudiante. Cabeza-Rodríguez [8] identificó que la eficiencia percibida de los asistentes virtuales es un predictor determinante de la satisfacción estudiantil en entornos en línea, con una valoración significativamente mayor en el género femenino. Por otro lado, Leiva-Guerrero et al. [9] validaron la calidad de la retroalimentación generada por IA (específicamente GPT-4o) bajo la estructura de la "Escalera de Wilson", demostrando que la herramienta puede ofrecer orientaciones claras y valorativas, aunque se detectó una tasa de "alucinación" (información no solicitada) del 15%, lo que subraya la necesidad de supervisión experta. Asimismo, Puertas y Cano [24] confirmaron que el uso de estrategias de prompting como Few-Shots mejora la precisión de la evaluación automática en comparación con Zero-Shot.

En conclusión, la evidencia empírica analizada demuestra que la Inteligencia Artificial ha dejado de ser una tecnología

emergente para convertirse en una herramienta didáctica con efectos medibles y positivos en el rendimiento académico, especialmente en matemáticas. Los datos confirman que, cuando la IA se implementa mediante diseños instruccionales planificados ya sea a través de modelos adaptativos o asistentes generativos bien configurados, no solo mejora la comprensión de conceptos complejos, sino que también eleva la motivación y la satisfacción del estudiante. Sin embargo, la presencia de errores técnicos y la variabilidad en la adopción tecnológica advierten que el éxito de estas herramientas no es automático, sino que depende intrínsecamente de la mediación docente y de una alfabetización digital crítica que permita al alumnado utilizar estos recursos como complementos para el razonamiento y no como sustitutos del pensamiento.

C. Impacto de la Inteligencia Artificial en la enseñanza y el aprendizaje de las matemáticas

TABLA IV:
IMPACTO DE HERRAMIENTAS DE IA EN EL APRENDIZAJE AUTODIRIGIDO Y LA ENSEÑANZA DE LAS MATEMÁTICAS

Referencia (Autor/Año)	Herramienta IA	Enfoque Principal	Impacto en la Enseñanza-Aprendizaje	Dimensión del Aprendizaje Autodirigido
Delgado Santin et al. [3]	Modelo adaptativo (basado en GPT-4 y diagnóstico de errores)	Razonamiento algebraico y personalización	Mejora del 33.1% en el rendimiento del grupo experimental; incremento de la percepción de autoeficacia.	Resolución de problemas y Motivación (aumento de la confianza matemática).
Plaza García et al. [5]	ChatGPT	Aprendizaje de funciones matemáticas (Dominio y Recorrido)	Diferencia significativa en calificaciones (7.75 vs 5.79 en grupo control); comprensión profunda de conceptos abstractos.	Resolución de problemas (comprensión conceptual) y Motivación (mayor interacción).
Leiva-Guerrero et al. [9]	GPT-4o (Estructurado con la Escalera de Wilson)	Retroalimentación formativa y evaluativa	Retroalimentación clara, útil e inmediata que activa conocimientos previos; los estudiantes valoran la rapidez, aunque critican la falta de "toque humano".	Autorregulación (reflexión sobre el propio proceso de aprendizaje y metacognición).
Guisheca Ayala et al. [15]	Plataformas de enseñanza respaldadas por IA	Enseñanza personalizada en educación básica	Aumento del 30% en la capacidad de resolución de problemas; reducción de la ansiedad matemática al permitir avanzar al propio ritmo.	Motivación (reducción de ansiedad) y Autonomía (progreso a ritmo propio).
Bonilla Marín et al. [4]	ChatGPT, Gemini, Copilot, GeoGebra	Cálculo multivariado e Ingeniería de Prompts	Mejora en la visualización de problemas y agilidad en procedimientos; permite al estudiante validar resultados y	Autonomía (aprendizaje a ritmo propio) y Resolución de problemas (verificación crítica).

			detectar "alucinaciones" de la IA.	
Ramírez-Medina et al. [21]	Plataformas digitales impulsadas por IA	Rendimiento académico en educación superior	Correlación alta entre el tipo de herramienta IA utilizada y el rendimiento; mejora en la organización del estudio.	Autorregulación (planificación académica y gestión del tiempo).
Cabeza-Rodríguez [8]	Asistente virtual basado en ChatGPT 3.5	Satisfacción y eficiencia en educación en línea	La eficiencia percibida del asistente es el factor más relevante para la satisfacción; facilita la resolución rápida de dudas frecuentes.	Autonomía (apoyo al aprendizaje independiente y reducción de dependencia del docente).
Patac y Patac [7]	ChatGPT	Gestión de la carga cognitiva	Simplificación de tareas complejas y reducción de la carga cognitiva extrínseca, liberando recursos mentales para el razonamiento.	Autorregulación (gestión del esfuerzo cognitivo) y Autonomía.

El análisis de las herramientas y estrategias pedagógicas sistematizadas en la Tabla 4 revela que la integración de la Inteligencia Artificial (IA) en la educación matemática genera un impacto significativo y multidimensional, transformando tanto el rendimiento académico cuantificable como la calidad de los procesos cognitivos y afectivos asociados al aprendizaje de esta disciplina.

Mejora del rendimiento académico y comprensión conceptual: La evidencia empírica indica que el uso de asistentes generativos y modelos adaptativos supera a los métodos de enseñanza tradicionales en temas de alta abstracción. En el nivel de bachillerato, Plaza García et al. [5] demostraron que la interacción con ChatGPT para el estudio de funciones matemáticas resultó en una mejora sustancial de las calificaciones, donde el grupo experimental alcanzó un promedio de 7.75 frente al 5.79 del grupo control. Este hallazgo se complementa con lo reportado por Delgado Santin et al. [3], quienes observaron una mejora del 33.1% en el rendimiento de estudiantes de secundaria al utilizar un modelo adaptativo para el razonamiento algebraico, facilitando la detección precisa de errores y la personalización de la ruta de aprendizaje.

Personalización, ritmo y reducción de la ansiedad matemática: Uno de los impactos más profundos de la IA es su capacidad para mitigar barreras emocionales que históricamente dificultan el aprendizaje matemático. Guisheca Ayala et al. [15] encontraron que las plataformas personalizadas en educación básica no solo incrementaron en un 30% la capacidad de resolución de problemas, sino que redujeron significativamente la ansiedad matemática al permitir que los estudiantes avanzaran a su propio ritmo. Esta adaptabilidad fomenta una mayor autonomía y motivación intrínseca, elementos críticos para sostener el esfuerzo cognitivo en tareas complejas.

Desarrollo de competencias en educación superior y el rol del *Prompt Engineering*: En contextos universitarios, específicamente en cursos de cálculo multivariado, Bonilla Marín et al. [4] identificaron que la IA Generativa (como ChatGPT, Gemini y Copilot) actúa como un catalizador para la visualización de problemas y la agilidad procedimental. Sin embargo, el estudio subraya que el impacto positivo depende de la alfabetización digital del estudiante: el éxito en el uso de estas herramientas está intrínsecamente ligado a la capacidad de formular instrucciones precisas (*prompt engineering*) y, crucialmente, a la habilidad para validar críticamente los resultados, dado el riesgo de "alucinaciones" o respuestas inexactas por parte de la IA.

Retroalimentación formativa y metacognición: Finalmente, la IA ha redefinido la retroalimentación, pasando de ser un proceso correctivo diferido a uno formativo e inmediato. Leiva-Guerrero et al. [9] validaron el uso de GPT-4o estructurado bajo la "Escalera de Wilson", demostrando que la IA puede ofrecer orientaciones claras y sugerencias constructivas que activan conocimientos previos. Esta inmediatez permite al estudiante reflexionar sobre su propio proceso de aprendizaje (autorregulación) en el momento exacto en que surge la duda, una dinámica difícil de replicar masivamente en aulas tradicionales.

En conclusión, la integración de la Inteligencia Artificial en la educación matemática ha trascendido su rol como simple herramienta de cálculo para convertirse en un tutor personalizado y dinámico. La evidencia sugiere que, cuando se implementa con una estrategia pedagógica clara, la IA no solo eleva los indicadores de rendimiento académico, sino que transforma la relación del estudiante con la matemática, disminuyendo la ansiedad y fomentando la autonomía. No obstante, este avance tecnológico plantea un nuevo imperativo educativo: la necesidad de formar a los estudiantes no solo en la resolución de problemas, sino en la validación crítica de las soluciones generadas por la máquina, asegurando que la tecnología sea un andamiaje para el pensamiento y no un sustituto del razonamiento lógico.

D. Evaluación de la calidad metodológica de estudios empíricos representativos

TABLA V:
EVALUACIÓN DE LA CALIDAD METODOLÓGICA DE ESTUDIOS
EMPÍRICOS REPRESENTATIVOS

Referencia	Diseño del Estudio	Tamaño de Muestra	Herramienta	Calidad	Justificación
Ramírez-Medina et al. [21]	Correlacional	n = 569	Estándar estadístico	Alta	Muestra censal robusta; alta fiabilidad de instrumentos ($\alpha > 0.80$).
Delgado Santin et al. [3]	Cuasi-experimental	n = 240	Estándar estadístico	Alta	Grupo control definido; pre/post-test riguroso.
Leiva-Guerrero et al. [9]	Validación (Delphi)	n = 144	CASP	Alta	Triangulación sólida y validación por expertos.
Cabeza-Rodríguez [8]	Caso (Mixto)	n = 391	CASP	Moderada	Muestra amplia, pero con muestreo no probabilístico por conveniencia.
Plaza García et al. [5]	Experimental	n = 28	Estándar estadístico	Moderada	Diseño riguroso, pero con tamaño de muestra pequeño que limita generalización.
Marcillo et al. [23]	Estudio de caso	n = 25	CASP	Baja	Muestra muy reducida; resultados sin diferencias estadísticamente significativas claras.

Evidencia de alta calidad metodológica: Los estudios catalogados con alta calidad metodológica proporcionan la evidencia estadística más contundente sobre los beneficios de la IAG, destacándose por sus muestras robustas y el control de variables. Delgado Santin et al. [3], mediante un diseño cuasi-experimental con una muestra amplia (n = 240), demostraron que el uso de un modelo adaptativo genera una mejora del 33.1% en las habilidades de razonamiento algebraico del grupo experimental, frente a un 10.2% del grupo control, logrando un tamaño del efecto grande. En esta misma línea de alto rigor cuantitativo, Ramírez-Medina et al. [21], a través de una evaluación correlacional censal sobre 569 estudiantes universitarios, evidenciaron relaciones positivas y estadísticamente significativas entre la frecuencia de uso de herramientas de IA y dimensiones clave del rendimiento académico, como la comprensión de contenidos y la planificación. Desde un enfoque cualitativo y de validación experta, Leiva-Guerrero et al. [9] confirmaron mediante el método Delphi (con 8 especialistas) que el uso de GPT-4o para estructurar retroalimentación bajo la "Escalera de Wilson" obtiene valoraciones sobresalientes, fomentando la metacognición al brindar respuestas inmediatas, claras y útiles.

Evidencia de calidad moderada y baja: Los estudios de calidad moderada respaldan los hallazgos positivos anteriores, pero advierten limitaciones en sus muestras que restringen su nivel de generalización. Por un lado, Cabeza-Rodríguez [8] observó que la eficiencia de los asistentes virtuales basados en ChatGPT impacta favorablemente en la satisfacción y el aprendizaje de una amplia muestra de estudiantes en línea (n = 391); sin embargo, el uso de un muestreo no probabilístico por conveniencia modera la solidez de estas conclusiones poblacionales. Igualmente, bajo un diseño experimental riguroso, Plaza García et al. [5] comprobaron que los estudiantes que utilizaron ChatGPT para el aprendizaje de funciones matemáticas aumentaron significativamente sus calificaciones frente al grupo de metodologías tradicionales; no obstante, el tamaño reducido de la muestra (n = 28) limitó la validez externa del hallazgo.

En contraste, los estudios evaluados con baja calidad metodológica, como el diseño de caso de Marcillo et al. [23] con una muestra muy reducida (n = 25), no lograron demostrar diferencias estadísticamente significativas (p = 0.2970) tras la implementación de aplicaciones tecnológicas. Esto evidencia que la simple exposición a herramientas digitales, si carece de un marco metodológico estructurado y una muestra representativa, no garantiza mejoras inmediatas en el aprendizaje.

En conclusión, la evaluación de la calidad metodológica demuestra que el impacto de la inteligencia artificial en la educación matemática es altamente positivo y transformador siempre que se fundamente en investigaciones con gran rigor estadístico y diseños instruccionales estructurados. Los estudios de calidad alta confirman de manera inequívoca que estas herramientas potencian el rendimiento, la comprensión abstracta y la autorregulación del estudiante mediante retroalimentación precisa y personalizada. Sin embargo, las inconsistencias detectadas en los estudios de menor nivel evidencian que el éxito de la tecnología no es inherente a su mero uso; requiere obligatoriamente de una planificación pedagógica adecuada, validación experta y control de variables para evitar sesgos y maximizar verdaderamente el desarrollo cognitivo en el aula.

IV. DISCUSIÓN

El análisis de la literatura científica publicada entre 2022 y 2026 revela que la integración de tutores virtuales basados en Inteligencia Artificial Generativa (IAG) marca un punto de inflexión en la enseñanza de las matemáticas universitarias. A diferencia de los Sistemas Tutores Inteligentes (STI) tradicionales, que operaban bajo reglas predefinidas, los modelos de lenguaje grande (LLM) como ChatGPT, Gemini o Copilot actúan como agentes conversacionales dinámicos capaces de ofrecer una personalización sin precedentes, aunque su implementación no está exenta de tensiones pedagógicas y éticas [19], [18].

Tendencias pedagógicas: De la instrucción a la mediación dialógica Una tendencia predominante identificada es el tránsito de un modelo de transmisión de conocimientos hacia uno de mediación dialógica y adaptativa. La IAG permite la creación de rutas de aprendizaje personalizadas que se ajustan al nivel cognitivo y ritmo del estudiante, facilitando la comprensión de conceptos abstractos como el cálculo multivariado o el razonamiento algebraico [3], [4]. En este contexto, el rol del docente evoluciona de ser la única fuente de validación a convertirse en un diseñador de experiencias y un "ingeniero de *prompts*", encargado de guiar la interacción entre el estudiante y la máquina para asegurar la precisión matemática y evitar el aprendizaje superficial [10].

Impacto en el aprendizaje autodirigido y la autonomía La evidencia sugiere que los tutores virtuales actúan como un andamiaje efectivo para el aprendizaje autorregulado. Herramientas como ChatGPT reducen la carga cognitiva y la ansiedad matemática al proporcionar un entorno libre de juicios donde el estudiante puede practicar repetidamente y recibir retroalimentación inmediata [7], [15]. Estudios recientes demuestran que la retroalimentación generada por IAG, cuando se estructura bajo modelos pedagógicos como la "Escalera de Wilson", no solo corrige errores, sino que fomenta la metacognición, permitiendo al estudiante reflexionar sobre sus propios procesos de resolución de problemas [9]. Sin embargo, esta autonomía tiene un matiz crítico: si no se gestiona adecuadamente, puede derivar en una "falsa autonomía" donde el estudiante delega el razonamiento a la IA, debilitando su agencia intelectual [10].

Desafíos críticos: Alucinaciones y dependencia tecnológica A pesar de los beneficios en el rendimiento académico reportados en estudios experimentales [5], la fiabilidad de la IAG sigue siendo un desafío mayor. La literatura documenta consistentemente el fenómeno de las "alucinaciones", donde la IA genera procedimientos matemáticos plausibles pero incorrectos o inventa datos [4], [10]. Este riesgo exige que los estudiantes desarrollen un pensamiento crítico robusto para validar las respuestas de la máquina, una competencia que paradójicamente requiere un dominio previo de la materia. Además, existe una preocupación latente sobre la integridad académica y la equidad digital, ya que el acceso desigual a versiones avanzadas de estas herramientas (como GPT-4o) podría exacerbar las brechas educativas existentes [17], [11].

V. CONCLUSIONES

La presente revisión sistemática permite concluir que los tutores virtuales basados en Inteligencia Artificial Generativa se han consolidado como herramientas disruptivas con la capacidad de democratizar la personalización del aprendizaje matemático. Su impacto positivo es evidente en la mejora del rendimiento académico, la reducción de barreras emocionales hacia las matemáticas y la optimización de la gestión del tiempo de estudio mediante retroalimentación instantánea y adaptativa.

Sin embargo, la efectividad de estas herramientas depende intrínsecamente de la alfabetización digital crítica tanto de docentes como de estudiantes. La IAG no debe visualizarse como un sustituto del profesor, sino como un "socio cognitivo" que requiere supervisión experta. Los hallazgos subrayan la urgencia de integrar formalmente la ingeniería de *prompts* y la ética de la IA en los currículos universitarios, transformando el error algorítmico en una oportunidad pedagógica para el análisis crítico.

Para el futuro de la enseñanza y aprendizaje de las matemáticas, el desafío no reside en la adopción tecnológica *per se*, sino en el diseño de estrategias didácticas híbridas que equilibren la eficiencia de la automatización con la profundidad del razonamiento humano, garantizando que la autonomía del estudiante se fortalezca a través de la tecnología y no a costa de ella.

La restricción al idioma español constituye la principal limitación de esta revisión. Los autores reconocen que, al excluir investigaciones publicadas en inglés o portugués, es altamente probable que se hayan omitido tendencias globales, debates teóricos emergentes y hallazgos tecnológicos recientes.

En consecuencia, las implicaciones y resultados de este estudio no son directamente generalizables a nivel mundial. Los hallazgos reflejan las realidades de la cultura educativa hispanohablante e iberoamericana, un contexto marcado por desafíos específicos de equidad digital, políticas institucionales en desarrollo y niveles heterogéneos de infraestructura tecnológica. Las futuras investigaciones deberán ampliar los criterios idiomáticos para contrastar estas dinámicas regionales con los avances del panorama científico global.

REFERENCIAS

- [1] E. M. Reyes Vallejo, S. J. Changuan Loor, H. P. Almache Soria, y G. A. Villamarín Peñafiel, "Algoritmos matemáticos para inteligencia artificial en la educación superior," *REMULCI*, vol. 3, 2025. <https://doi.org/10.59282/remulci.3.1.1129>
- [2] L. Guardia Ortiz, Z. Bekerman, y M. Zapata Ros, "Presentación del número especial "IA generativa, ChatGPT y Educación. Consecuencias para el Aprendizaje Inteligente y la Evaluación Educativa"," *RED. Revista de Educación a Distancia*, vol. 24, no. 78, 2024. <http://dx.doi.org/10.6018/red.609801>
- [3] V. del C. Delgado Santín, M. L. Intriago Cedeño, J. A. Intriago Moreira, C. E. González Ramírez, y P. del P. Tandayamo Vargas, "Aprendizaje Matemático personalizado mediante inteligencia artificial: un modelo adaptativo para el desarrollo del razonamiento algebraico en educación," *ASCE Magazine*, vol. 4, no. 3, pp. 311–332, 2025. <https://doi.org/10.70577/ASCE/311.332/2025>
- [4] A. Bonilla Marín, J. E. Márquez Díaz, L. G. Benavides Ramírez, y F. R. Gutiérrez Arévalo, "Inteligencia Artificial Generativa (IAG) en la educación matemática," en *EIEI ACOFI 2024*, 2024.
- [5] M. A. Plaza García, M. A. Mendoza Mosquera, R. Avello Martínez, y T. Tapia Bastidas, "Uso de ChatGPT para mejorar el aprendizaje de las funciones matemáticas en segundo de bachillerato," *Revista Conrado*, vol. 21, no. 102, e4257, 2025.
- [6] M. E. Torres Valladares, A. M. Abanto Gálvez, P. M. Villamizar Morales, y M. Ugaz Rivero, "Aprendizaje autorregulado en entornos educativos virtuales universitarios: una revisión sistemática (2018-2023)," *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, pp. 151-166, 2024.
- [7] L. P. Patac y A. V. Patac, "Using ChatGPT for academic support: Managing cognitive load and enhancing learning efficiency – A phenomenological approach," *Social Sciences and Humanities Open*, vol. 11, 2025.
- [8] M.-Á. Cabeza-Rodríguez, "Asistentes ChatGPT en educación superior en línea y satisfacción del alumnado: un caso de estudio," *RIED-Revista Iberoamericana de Educación a Distancia*, vol. 28, no. 2, pp. 9-38, 2025. <https://doi.org/10.5944/ried.28.2.43552>
- [9] M. V. Leiva-Guerrero, I. Araya Zamorano, R. Escobar Collins, y F. Silva Castro, "Retroalimentación de aprendizajes con inteligencia artificial generativa en estudiantes universitarios," *RIED-Revista Iberoamericana de Educación a Distancia*, vol. 29, no. 1, pp. 241-265, 2026.
- [10] L. F. Díaz Basurco, "El uso de aplicaciones de la Inteligencia Artificial Generativa en la enseñanza-aprendizaje de las Matemáticas a nivel universitario: aspectos favorables y desfavorables," Tesis de Licenciatura, Pontificia Universidad Católica del Perú, Repositorio Institucional PUCP, 2025.

- [11] A. F. Vera-Fernández y N. A. Sánchez-Goycochea, "Chatbots como Herramientas de Soporte en la Gestión y Decisiones Educativas. Una Revisión Sistemática," en *23rd LACCEI International Multi-Conference for Engineering, Education, and Technology*, 2025.
- [12] J. J. Gutiérrez-Castillo, R. Romero Tena, y A. León-Garrido, "Beneficios de la Inteligencia Artificial en el aprendizaje de los estudiantes universitarios: una revisión sistemática," *Educec, Revista Electrónica de Tecnología Educativa*, no. 91, pp. 185-206, 2025. <https://doi.org/10.21556/edutece.2025.91.3607>
- [13] E. R. Reyes-Moreno, J. A. Londoño-Gallego, I. C. Andrade-Martelo, H. F. Villar-Vega, y J. J. Castro-Maldonado, "ChatGPT en la educación: un enfoque bibliométrico de la integración de sistemas de Chatbots en los procesos educativos," *Revista Latinoamericana de Ingeniería y Aplicaciones (LATIA)*, 2023.
- [14] F. A. Fernández Otoyá, M. O. González Fernández, P. Quadros-Flores, y M. Sánchez García, "Posibilidades de la IA en el prácticum: visión estudiantil desde la literatura," *Revista Prácticum*, vol. 10, no. 1, pp. 112-127, 2025. <https://doi.org/10.24310/rep.10.1.2025.21952>
- [15] L. A. Guishca Ayala et al., "Integración de la inteligencia artificial en la enseñanza de matemáticas: un enfoque personalizado para mejorar el aprendizaje," *Ciencia Latina Revista Científica Multidisciplinar*, vol. 8, no. 5, pp. 818-839, 2024. https://doi.org/10.37811/cl_rcm.v8i5.14114
- [16] L. L. Maguiña Huerta y J. E. A. Padilla Caballero, "Innovaciones tecnológicas en el aprendizaje de matemáticas en educación básica: revisión sistemática," *Revista de Climatología*, vol. 25, 2025.
- [17] M. Reina-Parrado, P. Román-Graván, y C. Hervás-Gómez, "Revisión PRISMA sobre el Aprendizaje Automático en la educación: Retos sociales y oportunidades en la formación de ciudadanos del mañana," *European Public & Social Innovation Review*, vol. 11, pp. 1-26, 2026. <https://doi.org/10.31637/epsir-2026-1620>
- [18] M. González Murillo, N. Cosme Toribio, y G. J. Ortega Jalil, "Estrategias Didácticas Mediadas por ChatGPT en la Enseñanza Universitaria de la Matemática: Una Revisión de Alcance," *Revista Colón Ciencias, Tecnología y Negocios*, vol. 13, no. 1, pp. 130-149, 2026. <https://doi.org/10.48204/j.colonciencias.v13n1.a8546>
- [19] D. Rodríguez Cairo y Y. Ramírez Echavarría, "Tutores Inteligentes: mejora de la calidad de la educación superior a partir de la IA," *LatIA*, vol. 1, no. 8, 2023.
- [20] K. L. Reyes Contreras, P. M. Bejarano Álvarez, A. C. Contreras Litardo, y T. Chara de los Ríos, "Tutoría entre pares en el aprendizaje de las matemáticas," 2025.
- [21] R. Ramírez-Medina, J. Zapata-Peña, D. Aguilar-Chuquizuta, y E. Bastidas-Muñoz, "Plataformas digitales de aprendizaje impulsadas por inteligencia artificial y su desempeño en la educación superior," *International Journal of Computational Innovations, Intelligent Systems and AI*, vol. 2, no. 1, pp. 48-69, 2026.
- [22] A. Sorbara, "La inteligencia artificial y el aprendizaje," *Dilemas Contemporáneos: Educación, Política y Valores*, 2023.
- [23] F. Marcillo y F. Carrasco, "App Pyto: un caso de estudio en el aprendizaje de matemáticas discretas," *Revista Científica de Informática ENCRIPITAR*, vol. 8, no. 15, 2025.
- [24] E. Puertas Prats y M. E. Cano García, "¿Puede la inteligencia artificial proporcionar un feedback más sostenible?," *Digital Education Review*, no. 45, pp. 50-58, 2024.
- [25] Y. R. Calla Chumpisuca, V. R. Nomberto Bazán, B. Mendoza Palomino, R. B. Ortiz Aucapiña, y I. K. Rimascua Rodríguez, "Actitudes hacia la inteligencia artificial y su relación con la satisfacción académica: el rol mediador de la comodidad en su uso educativo," *Bordón, Revista de Pedagogía*, vol. 77, no. 4, pp. 117-137, 2025. <https://doi.org/10.13042/Bordon.2025.112199>
- [26] M. C. Pezzini, "Tutores inteligentes en la enseñanza: Una revisión y análisis en la educación secundaria," Trabajo final de especialización, Universidad Nacional de La Plata, 2022.